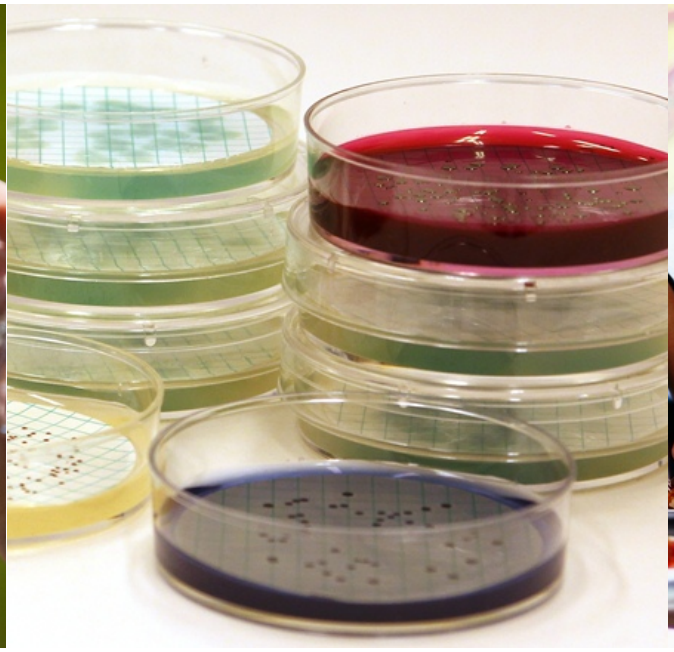
Proficiency Testing

PROTOCOL

Microbiology

Drinking water & Food


Livsmedelsverket
National Food Agency

NFA PT
Since 1981

A resolution on accreditation of the program was made 2004-12-10

SWEDAC ACKREDITERING
1457
ISO/IEC 17043

Fifth edition
2017-10-03

Corrected version (ver. 1.01)
2018-02-02

Editor-in-chief
Hans Lindmark, head of Biology department

Proficiency Testing	*E-mail*	*Telephone (operator)*
Microbiology	PT-micro@slv.se	+46 (0)18 17 55 00
	Internet	*Fax*
	www2.slv.se/absint	+46 (0)18 17 14 94

Contents

1. Introduction

Laboratories that perform analyses need to know that they obtain realistic results and to prove it to their clients in order to be trustworthy. They can obtain this knowledge by performing different kinds of controls. Internal controls may be used in order to see that no unexpected changes have occurred within the laboratory. Since microbiological analytical results are dependent on which analytical method is used, it is also important to be able to compare ones analytical results with those of other laboratories. One way to do this is to participate in interlaboratory comparative tests. These tests, when they are available, are compulsory for laboratories that aspire to become or are already accredited for their analyses. This is e.g. a requirement according to the standard EN ISO/IEC 17025 (1) where the name Proficiency Testing (PT) is used for these interlaboratory comparisons.

Proficiency testings are normally organised by a third party, independent of the laboratories and their clients. The evaluations are made by this party and the process is defined as an external assessment of the analytical competence of the laboratory.

The National Food Agency of Sweden provides such tests in the areas of food microbiology and drinking water microbiology and is accredited for this activity according to ISO/IEC 17043 (2). The tests are mainly directed to accredited laboratories within these analytical areas. The tests are also suitable for non-accredited laboratories, e.g. laboratories that perform controls in production lines of food or drinking water, but wish to be able to compare their analytical results with other participants.

The purpose of this protocol is to give to participants, and to other interested laboratories and parties, a description of the organisation of these microbiological proficiency testing activities and how some basic tasks are performed. In particular this protocol includes general parts that are not described elsewhere, including the production and handling of test material and the statistical processing of results. General information as well as specific information about the testing rounds is available on the program website *(see the front cover inner side)*.

2. Organisation

2.1. General information

Address: The National Food Agency
Box 622
SE-751 26 Uppsala
Sweden

Telephone: +46 (0)18 17 55 00

The National Food Agency is the central Swedish authority for food issues, including drinking water. Therefore, the National Food Agency organises microbiological proficiency testings, divided into one scheme for food and one scheme for drinking water.

E-mail for questions and opinions concerning the schemes: ***PT-micro@slv.se***

2.2. Information on the program website

2.2.1. Structure of the website

A general description of the schemes can be found on the website *(see the front cover inner side)*.

The website consists of two parts: one public and one for participants only where a user identity is required to log in.

2.2.2. The public section

This section contains pages that deal with:

- the proficiency testing activity in a general perspective, including the most recent edition of this document as a pdf file
- basic information regarding the drinking water and food schemes, e.g. which analytical parameters they comprise
- the content of each testing round along with the time schedule
- the conditions for participation
- the current prices
- addresses to contact the National Food Agency as provider
- a web form to register as a new participant in a scheme
- the login to participants' pages

2.2.3. The participant section accessible by login

This section comprises:

- one page to report the analytical results from the current round
- one page with a web form, to report details for each method in order to store them directly in our data base
- one page with the reported results and the calculated results from previous rounds that are presented 1) as preliminary shortly after finished recording of the results (to compare your reported results with the preliminary statistics), 2) as preliminary during processing of the results, when only the reported results and the preliminary median values are presented, 3) as final when a proficiency test has been fully evaluated and the laboratory results and the statistics correspond to those present in the final report *(as the web page is not as manually flexible as the written report, the written report overrules the web page when there might be a difference)*.

- one page for electronic registration/cancellation of participation in coming rounds
- one page with the reports and information to the participants (e.g. information letters) available as pdf-files.

2.3. Areas of responsibility

For each scheme, there is a *scheme coordinator* who has the overall responsibility but also specific responsibilities for e.g. planning, correspondence, processing of results and reports. For both schemes together, there is a *laboratory staff* responsible for the manufacture, quality checks and storage of the test material, before dispatch. There is also an *administrative support* responsible for the address register, the contacts regarding invoices, and laboratory contacts regarding e.g. participation, as well as dispatch of the testing material.

The overall responsibility to maintain the quality and the quality system of the program to a necessary level is assigned to the team manager together with the head of the Biology department.

The head of the Biology department is also the editor-in-chief of the final reports and this scheme protocol.

2.4. Advisory group

2.4.1. Composition and meetings

For each scheme, there is an advisory group composed of representatives from the countries with the largest numbers of participating laboratories, normally the Nordic countries. The two advisory groups meet in Uppsala every second or third year, both together as one group discussing issues common to the two schemes and parted discussing issues specific to the individual schemes.

2.4.2. Duties

The members can come from different organisations and countries, and represent both these and themselves. Their role is mainly advisory, with opinions on e.g. analytical parameters, frequency, costs, accepted methods and the content of the reports. Major changes within a scheme should be sanctioned by the corresponding advisory group.

2.4.3. Members

A list of the group members is available as a pdf-document under the heading General scheme information under the tab Info & Reports in the participant section of the website.

2.5. Accreditation

The National Food Agency is since December 2004 accredited for arranging microbiological proficiency testing. The accreditation has been approved by

Swedac, and is currently performed according to the standard EN ISO/IEC 17043:2010 (2).

3. The two schemes

3.1. The Drinking water scheme

The Drinking water scheme has been ongoing with participants from several countries since 1992.

The scheme comprises 2-4 test items twice a year. It includes about 10 quantitative analytical parameters of bacteria, moulds and yeasts with a focus on indicator organisms, including some that may cause illness.

Some parameters are part of both testing rounds. The remaining parameters are allocated between the two rounds. The parameters are presented on the page General Information, Drinking water scheme, on the website.

3.2. The Food scheme

The Food scheme has been ongoing with participants from several countries since 1988.

The scheme comprises 3-4 test items 3 times a year. It contains about 25 different quantitative and qualitative analytical parameters – bacteria, moulds and yeasts – including the analyses of pathogenic bacteria normally searched for in food.

Some parameters are part of all testing rounds. The remaining parameters are allocated between the three rounds. The parameters are presented on the page General Information, Food scheme, on the website.

3.3. Time schedule and analytical parameters for a testing round

Time schedule and analytical parameters for current testing rounds are found on the page PT rounds on the website.

As providers, we will do our best to respect the time schedules. In case of unforeseen events, we however, reserve ourselves the right to postpone particular round one or two months or ultimately to cancel it. The participants will be informed about such actions prior to the originally scheduled dispatch date of a round.

4. The test material

4.1. Type of material

Natural samples, or for the particular purpose manufactured test items, may be used in proficiency testing. A variant in between, where a cultured test organism is added to a natural or artificial product (”spiking” of a food or drinking water), may also be used at times.

The National Food Agency has chosen to use manufactured test items for the microbiological proficiency testing. These simulate food or drinking water samples containing mixtures of organisms, with each test material having a certain purpose.

Depending on the parameters of the round, the mixtures may include both bacteria and fungi (moulds and/or yeasts). Some mixtures include pathogenic bacteria, while others contain organisms possessing specific indicative properties only.

No test material containing protozoa or virus is manufactured.

The ready-to-use test material consists of 0.5 ml freeze-dried serum broth with different microbial mixtures in a 2 ml glass vial. The material is manufactured according to the description by Peterz and Steneryd (3). The sample for testing (or homogenate concerning food) is obtained after dissolving the material in a specific volume of suitable liquid.

Freeze-dried test material without matrix in vials has a number of advantages compared to natural samples or ”spiked” samples, but also some disadvantages.

4.1.1. Advantages

+ The vials do not require much space, which facilitates storage, package and transport.
+ The vials are fairly robust and are therefore easy to transport.
+ Several organism groups (analytic parameters) may be tested in parallel by the same test item.
+ The test material has a good stability and may be used in more than one round without appreciable change of microbial concentrations
+ It is possible to manufacture a large number of identical sub-samples where the organisms are Poisson distributed at lower concentrations or log-normal distributed at higher concentrations.
+ Expenses are kept low by having a standardised and relatively simple procedure of manufacture.
+ Dissolved test material may be used for “spiking” of natural matrices.

4.1.2. Disadvantages

– The freeze-drying procedure and machinery need to be sufficiently stable to ensure that the process can be repeated.

- The organisms require a protective substance, cryoprotectant, which helps them survive the freeze-drying.
- The material has to be dissolved in liquid, which requires a certain amount of work and can induce mistakes.
- The ingredients of the material may cause some foaming when it is dissolved. Therefore, the sample is likely to be distinguished from real samples, which means that the analyst may know it is a control sample.
- For food analyses, natural matrices are absent during preparation of the test material, since the dissolved test material itself mimics a ready-to-use homogenate.

4.2. Production and control

4.2.1. Composition

Freeze-dried cultures of each microorganism used in manufactured mixtures are stored at −70 °C in a collection at the National Food Agency. All strains are identified by specific SLV numbers. They have either been isolated from food and water samples or been bought from established culture collections. Bacterial strains are characterized internally by API system or by other means at external culture collections like ATCC (American Type Culture Collection) and CCUG (Culture Collection University of Gothenburg). The characterization of yeasts and moulds, *Salmonella* strain and *E. coli* O157 strains is performed externally by, the Centralbureau vor Schimmelcultures (CBS-KNAW Collection, The Netherlands), National Veterinary Institute (SVA, Sweden) and the Public Health Agency of Sweden, respectively.

ISB (Inositol Serum Broth) is used as cryoprotectant to protect the organisms when being freeze-dried. ISB is composed of sterile-filtered horse serum mixed with inositol and a small amount of nutrient broth. After washing and diluting organism cultures from nutrient rich medium (see below), the component SPG (Saccharose Phosphate Glutamate), which also possesses organism protective characteristics in the freeze-drying process, is added. Peptone water or a potassium phosphate buffer containing $MgSO_4$ (see e.g. ISO 8199:2005; 4) is used when diluting and washing. The same solutions can be used when dissolving ready-to-use material before analysis.

4.2.2. Manufacture and storage

Each organism is cultured individually in suitable nutrient medium, usually TSB (Tryptone Soy Broth) or BHI (Brain Heart Infusion) Broth/Agar for bacteria, and MEA (Malt Extract Agar) for fungi. The cultures are used after a predetermined time period of growth, after which the nutritious medium is washed or diluted off. The test mixture of organisms is obtained by transferring predetermined volumes of each diluted/washed culture (or spore suspension for moulds) to a specific volume of ice-cold ISB.

The test mixture is kept stirred on ice while aliquots (0.5 ml) are transferred into sterile glass vials placed on special racks. The vials are then freeze-dried and closed under vacuum according to specific procedures.

The produced test material is transferred to a freezer and tested for its organism content shortly after the manufacture. The vials from accepted mixtures are checked for vacuum and are sealed with aluminium caps. The test material is thereafter stored at ordinary freezer temperature (between –18 and –24 °C) or alternatively at –55 °C until used. The labelling is made in connection with the dispatch.

4.2.3. Quality controls

4.2.3.1. Purity of the strains

The purity of all strains included in the microbial mixtures is controlled by growth on unselective medium directly from the strains collection. In addition, the purity of each obtained culture is checked by streaking an aliquot on unselective medium. If there is any doubt in the purity of a strain, it is excluded from the mixture and the production is usually stopped.

4.2.3.2. Amount of test material

While dispensing the test mixtures into vials the dispensed volume is monitored by weighing. The total range and measure of dispersion (coefficient of variation) are calculated from the weights. The range may be at most 0.015 g between the dispensed amounts, which corresponds to 3.0 % of the average dispensed amount of material (the target volume 0.5 ml). The weight monitoring is done in parallel to the dispensing, in order to be able to adjust or interrupt the dispensing process, if the weights vary too much.

4.2.3.3. Concentration determination

Concentrations of the included organisms are determined after freeze-drying by using one or several vials of the test material. For new strains, the concentrations in a mixture are often determined before freeze-drying as well. When analyses are performed both before and after freeze-drying, the die off of the individual strains by the freeze-drying process can be determined (the reduction factor).

4.2.3.4. Homogeneity

Before the test material is used, the homogeneity must be checked and accepted. In this first test of homogeneity, duplicate analyses of 10 vials from the whole filling process (stratified sampling) are made with appropriate methods for the parameters to be tested. For the test material to be homogenous, certain criteria regarding variation within and between vials (see below clause 8.2 Determination of concentration and homogeneity) must be fulfilled.

4.2.3.5. Vacuum test

An inert environment is necessary in order to maintain the viability and stable concentrations of the test material. To ensure long-time stability of the material, the freeze-dried material therefore needs to stay under vacuum after the vials are

sealed and capped. Each individual vial is therefore tested for vacuum before storage or delivery, and vials without vacuum are discarded. Normally, very few of the newly produced vials need to be discarded.

4.2.3.6. Stability

The stability of many of the organisms included in the test materials has been investigated for several years in different freeze-dried mixtures. Each mixture is however not tested in this sense. A renewed concentration and homogeneity check is done only when used after a longer period of storage (see 8.3). Knowledge of long-term stability of the test material is based on similar material (reference material) that has been manufactured in the same manner, stored for at least 2 years and tested regularly. When stored frozen at low temperature, −55 °C most tested bacteria and fungi have shown stability for at least this period of time. Gram negative bacteria however, tend to decrease somewhat in colony recovery with time, while Gram positive bacteria and fungal spores are generally unaffected.

Under storage at approximately 25 °C, no noticeable negative effects on the stability of the test material has been observed for at least 3 weeks, except for the bacterial genus *Campylobacter*. At 44 °C during the same time, the recovery is somewhat affected, especially for Gram-negative bacteria. This tendency is even more noticeable after storage at 60 °C.

Considering these observations, the following guidelines apply:

- If stored at normal freezer temperature (–18 to –24 °C) the content is stable for at least one year, with the exception for Gram negative bacteria as described above.
- If stored in a refrigerator (5±3 °C), the content is stable for at least a few months, and usually considerably longer.
- If stored at room temperature, the content is stable for at least one month, and usually longer.

4.3. Hazards

4.3.1. Risk of infection

Microorganisms of the risk group 1 and 2 according to the classification of the Swedish Work Environment Authority (5) are used in the schemes. The former Swedish Institute for Infectious Disease Control (today the Public Health Agency of Sweden) made a risk assessment concerning the risks of infection by the organisms used in the schemes based on the way they are handled (6).). In their assessment they concluded that the risk for illness even after consumption of the contents of one vial is considered extremely small. Based on these statements, the National Food Agency made a stipulation regarding handling and transport (7; see further under 4.4.3.).

4.3.2. Environmental danger

The test material itself consists of horse serum, inositol and microorganisms, and usually also contains nutrient broth, SPG, peptone, potassium phosphate, and magnesium sulphate. The container consists of glass, rubber, aluminium and has a paper label. Since the container with material does not include any specifically classed or in any other way potentially dangerous chemical compound, it may be discarded in the common waste management after the microorganisms have been rendered harmless by killing. See below under 4.5.4 Destruction of the material.

4.4. Labelling and transport of test items

4.4.1. Randomisation of test vials to laboratories

4.4.1.1. Numbering of vial labels

Before dispatch of vials to the laboratories participating in a proficiency test, the vial labels are numbered. The numbering is made automatically and is saved in our data base. The numbers consist of the specific and confidential laboratory number followed by a hyphen and a digit which is connected to a specific test mixture. This digit is randomly picked for each laboratory for each of the different mixtures A, B, C etc. The link between digit and mixture is available from the data base.

4.4.1.2. Selection of vials for a laboratory

Before packaging, the vials of each mixture are carefully shuffled. For each laboratory, a vial is randomly picked and labelled with a number specific to the laboratory and the test mixture, as described above. This activity is carried out on a separate work bench for each test mixture in order to avoid errors.

4.4.2. Package of test vials

The individually labelled vials are placed in a transportation tube containing a shock- and liquid-absorbing material. The transportation tube is put in a protective envelope or in a safety jar that is put in a cardboard box. The package is marked with an address label and, when necessary, a customs declaration and/or other markings necessary for transport. Enclosed are instructions for the proficiency testing and a safety data sheet describing the content of the vials and that it can be regarded as a freeze-dried artificial food sample. The safety data sheet also describes how the material should be stored and destroyed as well as measures to take in case of leakage and physical contact with it.

4.4.3. Transport of test vials

Based on the risk assessment of the Swedish Public Health Agency (6) and the stipulation made by the National Food Agency (7), the test vials are packed as described previously, and sent via ordinary postal means. Padded envelopes are normally used, and a tracking service is used if needed. A courier service can also be used if requested by the laboratory.

4.5. Recommended handling of the test material upon delivery

4.5.1. Storage in connection to dispatch of test material

The material is kept cooled before package and transport (see 4.2.2).

During packaging and transport, the material is kept at ambient temperature.

When the test material is to be used shortly upon delivery, as in the case of proficiency tests, the needs of a long shelf life is relatively small. Storage in refrigerator is therefore satisfactory when the material is received. Storage at room temperature for a couple of weeks is also in general not critical. The test material should however always be kept in the dark.

The laboratories are always requested to keep the received material in darkness and in refrigerator or freezer until use.

4.5.2. Preparation of samples

The preparation is clearly described to the participating laboratory in the enclosed instructions, which include pictures with explanatory text. The procedure states that the test material should be transferred to a measured volume (e.g. 250 or 800 ml) of diluent. The suspension with the test material should then be carefully blended in order to obtain the sample ready for analysis.

4.5.3. Stability of the prepared samples

Once the freeze-dried material has been re-suspended, the concentrations cannot be presumed to be stable for longer than about an hour, not even after cooling. The prepared sample should therefore be used for analysis within one hour.

Some bacterial spores can however be maintained for a considerably longer period of time, e.g. several weeks for *Clostridium perfringens*.

4.5.4. Destruction of test material

4.5.4.1. Unopened vials containing test material

Before the material can be discarded in a regular waste bin, the microorganisms need to be killed. This can be done e.g. by autoclaving at 121 °C, long enough (e.g. 50 minutes) for all material to reach that temperature. An alternative is to hand in the vials containing the test material to a facility specialized in the destruction of infectious material.

4.5.4.2. Opened and used vials

A glass vial that contains/has contained test material, and the rubber stopper may suitably be discarded in containers for infectious material which are to be destroyed by a special facility. The aluminium cap can be discarded along with metal waste.

4.5.4.3. Remains of prepared sample

Remains of the prepared sample should be autoclaved at 121 °C for at least 15 minutes, or treated in any other way ensuring the destruction of the remaining microorganisms, before they are discarded.

5. Instructions for a particular testing round

5.1. Time schedule

Dispatch of test material and instructions is made 1–3 weeks ahead of the starting date for analyses in a testing round. The included written instructions state the starting date of the round and the final day to report the analytic results. These dates, as well as other dates for the respective round, are also stated on the website.

5.2. Analyses

The analytic parameters included in each testing round can be found on the website but are also stated in the sent instructions.

5.3. Other information

The instructions that are enclosed with the test material also contain information regarding:

- preparation of samples
- particular conditions concerning the different analyses, such as the dilutions or volumes that should be tested or the methods that can be used
- the reporting of method information, for each analysis in the participants pages of the website (this is mandatory to be able to report results).
- the reporting of the analytical results on the web form in the participants pages of the website, (a printed form with the analytical parameters and units is enclosed with the instructions, intended as back-up if there should be a problem with the web form)
- the use of the method information to group the analytical results per method in the final reports.

6. A testing round

6.1. Participant activities

6.1.1. Instructions and analyses

The participating laboratories are presumed to read the whole instruction and thus handle the test material according to the recommendations. They should however also, as far as possible, perform the analyses in the same manner as for routine samples, taking into account the restrictions or addenda stated in the instructions.

6.1.2. Reporting method information

Compulsory method information for an analytical parameter must be entered in order for the laboratory to be able to register analytical results. Method information can only be reported on the website and applies until further notice. It can be entered and adjusted any time after logging in, which means also after the last reporting date for analytic results and between testing rounds. The method details used to obtain the reported analytical results should be the ones reported.

Whenever there is a need to distinguish method differences in the reports, the method information details will be used to group the results. They will normally be used in the way they are reported in the data base at the closing date for results reporting. Although significant statistical differences may be difficult to prove, trends and possible disparities will be discussed, to help the interpretation of the varying results obtained by the laboratories. All possible groups of a parameter will not be discussed each time.

6.1.3. Reporting analytical results

The results are expressed differently in the drinking water and the food scheme. Therefore, the way to report the analytical results differs between the schemes, but is thoroughly described in the respective instructions.

The results of a participating laboratory have to be reported in time to be included and processed in the final report of a proficiency test. Before the deadline the reported results can be changed as many times as desired.

6.2. Modifying a reported result

6.2.1. The last day for reporting

It is the responsibility of the participating laboratory to report method details and results within stated time period.

However, laboratories are normally reminded by e-mail some days before the deadline that the closing date approaches. In this e-mail, the closing date is also explicitly stated.

6.2.2. Corrections

Only adjustments due to technical reasons (computer errors) are accepted, as well as result reporting errors made by us as organiser. If the error is caused by a participant claiming our instructions are imprecise or difficult to understand, corrections may be accepted after individual consideration.

Results errors due to mistakes made by the participating laboratory are not accepted and hence not corrected. Such errors are e.g. mistakes made when entering results, results reported for wrong sample, wrong analysis, wrong dilution, errors caused by calculation mistakes and results reported in any other way than described in the instruction.

On the other hand, method information can be corrected anytime by the participant, also in between testing rounds.

6.3. Possible sources of errors in a testing round

6.3.1. Registration/cancellation

In order to minimise the risk that laboratories in our data base that wish to participate should not receive test material, or the other way around, the laboratories are asked to make registrations and cancellations themselves, via a page for participants on the website It is also possible to sign in as subscriber for all or specific series of rounds, which means that the laboratory is registered for upcoming testing rounds in these series until it changes its status.

6.3.2. The test material

During manufacture, the amount of test material is systematically checked by sampling during filling of the vials. After freeze-drying, all vials are checked for vacuum, which is necessary for the survival of the organisms. Random samples (ca 10% of all vials) are checked once more for vacuum shortly before dispatch.

6.3.3. Dispatch and transport

Mix-up of samples among different laboratories is a possible risk at dispatch. As described in chapter 4.4.1, the vials are packaged in such a way that this should be avoided.

Addresses are continuously being updated as changes are received, in order to minimise address related mistakes.

In theory, the testing material could be damaged during transport if it is subjected to very high temperatures or strong x-rays. So far, these risks do not seem to have posed any evident problems, not even when material is transported very long distances to warmer countries. However, see clause 4.5.1. Information from “Posten” (the Swedish main messaging and logistic service) regarding domestic and international goods at Arlanda (the international airport of Stockholm) states that very low doses of x-rays are used (<1/100 of the dose for dental x-rays). Since no general negative effect from transportation has been noticed, it seems

likely that the doses of x-rays utilized are harmless for the freeze-dried material transported.

6.3.4. Incorrect reporting of analytical results and method information

The reporting of the analytic results is normally made by the participants themselves via a web form on a participant page on the website. The results are saved in the data base just as the participant entered them. There is no processing in between that can cause incorrect registration. If a participating laboratory enters an incorrect result, it is not regarded as a mistake that can be corrected, but is part of the proficiency test (see 6.2.2).

In special cases, as when there is trouble with the website or Internet connection, a participant can send the results to the organiser for entering to the database. The laboratory will then obtain an e-mail with the reported results attached. The results should be checked. Comments on typing errors etc. for correction must be sent to us within the accepted reporting time period.

The web form intended for method description is always available for changes and additions. It means that the participants themselves make necessary adjustments in that web form.

6.3.5. Inaccuracies in the final report

Should a significant error be found in the final report, the participants will be informed by e-mail. The report is adjusted and a new version is published on the website for reports and information. Less substantial inaccuracies can be corrected directly in the e-mail without publishing a new version of the report.

7. Follow-up of analyses

The National Food Agency does not request the participating laboratories to follow up the obtained results and to take suitable measures. Such demands can only be made by the laboratory itself or by a third party to which the laboratory is subordinated, e.g. an accreditation body. They can demand that the laboratory keeps a certain quality level and takes measures when the quality is questioned. How the follow-up should be done is therefore designed by the laboratory itself or in cooperation with the third party.

On the contrary, the National Food Agency does not take any responsibility for how or if the follow-up is done.

Z-scores (see below under Statistics) are given in an annex of the reports, as a means for the laboratories to evaluate their analyses. Z-scores are a good tool to evaluate an analytical parameter over several rounds, e.g. by a control chart.

As a provider, we also facilitate the follow-up by delivering extra vials of the test material to laboratories that ask for it – as long as stocks last. Each laboratory may receive one extra vial per test mixture free of charge. In order to receive a vial

from a mixture, the laboratory should in principle be able to state which analytical parameter they failed to analyse in that mixture.

8. Statistics and reporting

8.1. General

The statistical processing in the proficiency testing of microbiological drinking water and food laboratories includes the following.

- Checking the amount, concentrations, homogeneity and stability of the test material.
- Transforming the colony counts (CFU) before calculations, in order to obtain a normal distribution and a uniform variance within the range of results for the respective analysis. In the Food scheme, *log_{10} transformation* is made and in the Drinking water scheme, *square root transformation* is made.
- Identification of deviating analytical results as false positive and false negative results, as well as low and high outliers.
- Compilation of all results reported by the participating laboratories in a table, along with summary statistics (deviating results excluded) and the number of deviating results per analytical parameter.
- Visualisation of the results for each relevant quantitative analysis in a histogram for respective test mixture.
- Showing statistical evaluation and/or result distribution per sample for groups when different methods are used for the analysis of a parameter.
- Visualisation of the standardised analytical results (z-scores) of each laboratory in an individual box plot.
- Marking of the outliers and false results in the table containing all the analytical results and reporting the number of these deviating results for each laboratory below the respective box plot.

Results that are obviously erroneous (e.g. undoubtedly false), based on the knowledge of the test material, are excluded without any statistical test before the identification of outliers is made.

8.2. Determination of concentration and homogeneity

8.2.1. Matters common to both schemes

8.2.1.1. General approach

Accredited analytical methods are used when determining concentration and homogeneity of a test material. Methods that are not accredited may be used when a new parameter is tested or under certain special cases. This is noted when done.

Concentrations of the various organisms in a test material are determined in part to check that the material possesses all the qualities wanted, and in part to have reference values when evaluating the laboratory results. The determination of homogeneity is made based on the organism concentrations in vials from a stratified design (directly after manufacture), or by randomly selected vials (subsequent stability tests).

The homogeneity of the freeze-dried test material is normally checked before dispatch to the participating laboratories. Packaging is not considered to interfere with the homogeneity. Aliquots from several vials are consecutively examined by the same person. The relative variation, both between and within vials depends on which parameter is analysed but also on the concentration of the organism. A larger variation between vials in the homogeneity test usually implies a larger variation between the results of the participants as well. This is compensated for by the fact that the standard deviation for calculation of z-scores (see 8.6.6) is not fixed, but is a robust measure based on the results obtained by the participants.

The decisive criterion for homogeneity is in principle the same in the two programs. The tests of homogeneity are to some extent based on what is stated in international protocols (8, 9). These protocols are mainly worked out for quantitative chemical analyses and can therefore in some aspects not be strictly followed within microbiology. In the more recent editions of these protocols (10, 11), determination of homogeneity is treated somewhat differently. This latter procedure has not been considered applicable for the microbiological activity described here, and has therefore not been used. Even the previous way to determine homogeneity is used here only as a guideline during the control and is completed by other measures more fit for the purpose.

8.2.1.2. Specific calculations for the evaluation of homogeneity

Exceptionally, upon homogeneity test, one or both values from the duplicate analysis of a single vial can be deviant from the rest of the values obtained for the other vials leading to exceedance of the homogeneity criteria limits. In such a case, we reserve the right to recalculate the homogeneity criteria after excluding these values. This will happen if it is reasonable to think that the divergent results are due to another reason than non-homogeneity of the mixture, e.g. pipetting error or wrong volume analysed. If the new calculation shows homogeneity, the mixture will be approved.

1 – ANOVA

One-way analysis of variance is performed on the results from the 10 vials with duplicate determinations (8, also described in reference 19), since ANOVA is often used in proficiency testing when results have a normal distribution, e.g. like in chemistry. The analysis is carried out with square root transformed results for drinking water and $\log_{10}$ transformed results for food in order to obtain uniform variances and as good normal distribution as possible. An F-test is made to see that the dispersion between vials is not markedly larger than within the vials.

2 – ”Index of dispersion” – check of randomness

As a complement to the analysis of variance a test more suitable for microbiology is used. The test "Index of dispersion" is used to check that the obtained analytical results both within vials (10 duplicate analyses) and between vials (10 vials) do not differ markedly from what would be expected based on the appropriate Poisson distributions (13, 14). In these tests are used the original *non-transformed colony counts* from the sample volume chosen for the analysis of the parameter. The test is dependent on concentrations in the sense that it is easier to obtain acceptance for randomness, i.e. no contradiction of the Poisson distribution, at low colony counts compared to high.

3 – Test of reproducibility

The test (17) is independent of the others and of the microorganism concentration. It is therefore a necessary complement to the Index of dispersion. This test is performed on the 10 average values or sums (both gives the same results) from the two results obtained for each vial expressed in $\log_{10}$ units, to normalize the results in a general way.

8.2.2. The Drinking water scheme

8.2.2.1. Prerequisites and assumptions

Quantitative analytical results, concentration determinations as well as the participant's results are obtained by manual counting of colonies. A result may be derived from one or more different volumes of a sample (dissolved test material) but is in general converted into a pre-set analytical volume of the sample. Colony counts from different volumes of the sample are generally considered to be Poisson distributed. Since no dilution is usually made, the Poisson distribution is regarded to be valid, approximately, even for the converted results. It is however strictly valid only for those test volumes that are actually investigated. In a Poisson distribution, the variance is numerically equal to the average number of colonies.

As a consequence of assuming a valid Poisson distribution, based on theory (14), the square root transformation is used in order to approximate a normal distribution. Initial studies, made when the scheme was newly started, also in most cases showed that even the results obtained by the participants, proved to have best normal distribution when they were square root transformed, compared to log transformed or not transformed at all.

The analytical results are used square root transformed in determinations of concentrations, in analyses of variance to evaluate the homogeneity but also in detection of outliers, where a reasonable normal distribution is a prerequisite. In tests based on Poisson distributed results, the results are used without any transformation.

Depending on the testing round and the analytical parameters, the number of quantitative analytical results varies from approximately 40 to 100.

8.2.2.2. Initial check of concentration

As soon as possible after the freeze-drying of a mixture, 5 vials, randomly picked at different stages of the filling process (beginning, middle and end), are analysed. The content of each vial is dissolved in a fixed volume of diluent. Single analyses are performed from various sample volumes. This concentration check is used to get an early indication on whether the mixture is acceptable and assumed to be homogeneous with respect to the various organisms included. It is also used to decide which sample volume should be used in the final determination of concentration and homogeneity.

8.2.2.3. Final determination of concentration and homogeneity test

Before dispatch of the samples, 10 vials per mixture from the various stages of dispensation process (stratified sampling) are tested with respect to homogeneity and concentrations. Preparation of test material is done as for the initial concentration check. For each mixture, duplicate analyses are performed on the 10 vials during one and the same day. The sub-samples from the 10 vials cannot be analysed in a random order, since unwanted changes as regards the concentration of organisms can occur in the sample suspension if too much time elapses before analysis. The two sub-samples prepared from the same vial are therefore analysed consecutively, which may lead to under dispersion. Only one specified sample volume, based on the initial concentration check, is used for each parameter. The analyses are performed under repeatability conditions. A global average and a coefficient of variation (CV; see 8.6.7) are calculated from the 10 mean values (each from 2 determinations) per analysis.

8.2.2.4. Criteria for homogeneity

Results from the three tests described above complement each other and give a complete picture of the outcome. However they must be interpreted with care and caution, considering the results from previous similar mixtures. The calculations from ANOVA and "Index of dispersion" within vials are used to understand the contribution of randomness in the outcome of the tests. Two empirically verified indicative limits from the other tests are used as decision makers.

The actual criterion for homogeneity is that the values obtained for the test of reproducibility (T) and the test "Index of dispersion" between vials (I_2) *should not simultaneously exceed the set limits*, i.e. 2 both for T and I_2.

For production of test items before 2015, the guideline for homogeneity to be acceptable was that the coefficient of variation (CV, see 8.6.7) should not exceed 25% when the average content was at least 10 CFU per analysed sample volume. With a lower CFU average (<10 CFU; often poor normal distribution even after transformation), a CV higher than 25% was accepted if the distribution of the colony counts was as expected, based on the interpretation of the test "Index of dispersion" (with 95% confidence).

8.2.3. The Food scheme

8.2.3.1. Prerequisites and assumptions

Quantitative analytical results – concentration determinations as well as participants' results – are obtained by manual counting of colonies. The counts are thereafter in general converted to a pre-set analytical volume, taking into account

dilutions that may have been made. Colony counts obtained from one and the same dilution in a series are in general considered to be Poisson distributed. Since counts from several dilutions are often recorded, and since conversion is made back to a specified volume or weight of the original sample, results cannot fully be considered to strictly follow a Poisson distribution. The variance increases more than what is assumed in the Poisson distribution and therefore a log normal distribution is used as an approximation. The usage of common logarithms (log_{10}) is also the practice when it comes to microbiological food analyses.

The final calculated analytical results are therefore expressed as log_{10} results and are considered more or less normally distributed. These log-transformed results are used when calculations of concentrations, assessment of homogeneity and detection of outliers are made.

Depending on the testing round and the analytical parameter, the number of quantitative analytical results varies from approximately 40 to about 200 (in some cases also <40).

8.2.3.2. Initial check of concentrations

After the freeze-drying of a mixture, one vial is checked with respect to organism composition and concentrations. The content of the vial is dissolved in a fixed volume of diluent, making what is regarded as the zero dilution. Single analyses are made from various dilutions. The concentration check is used in order to obtain an indication on whether the test mixture is acceptable regarding the different included organisms and also to decide which dilution should be used for each analysis in the final determination of concentration and check of homogeneity.

8.2.3.3. Final determination of concentration and homogeneity test, quantitative analysis

Before dispatch of testing material, the same person analyse 10 vials from the various stages of dispensation process (stratified sampling) on the same occasion for a final determination of concentrations and test of the homogeneity of the test items. Preparation of the test material is done as for the initial concentration check. The sample is analysed either manually or with a spiral plating instrument. Two dilution series are made from each vial. For each analytical parameter one Petri dish per series is used from the most suitable dilution based on the initial concentration check. The two dilution series from each of the 10 vials cannot be analysed in a random order, since unwanted changes in the organisms concentrations can occur in the dissolved and diluted sample if too much time elapses before analysis. The two dilution series prepared from the same vial are therefore analysed consecutively, which may lead to an under dispersion. The analyses are performed under repeatability conditions. An average result is calculated for each vial, derived from the two Petri dishes. The final concentration is calculated as the mean value of these 10 average results, *log_{10} transformed*, after having been converted to the original volume taking the dilution into

account. The standard deviation for these 10 results, expressed in $\log_{10}$, is also calculated.

8.2.3.4. Final determination of concentration and homogeneity test, qualitative analysis

For determination of concentrations in qualitative analyses (e.g. *Salmonella*), separate vials containing a pure culture, special vials, are freeze-dried at the same time as the vials containing the mixed culture. The volumes of organism cultures are such that the same concentration can be expected in the special vials and in the mixed culture. The concentration is measured by analysing 10 special vials, from which duplicate determinations are made. Calculation of the mean value and the standard deviation is made as for the quantitative analyses, as well as the assessment of homogeneity. The mean value is considered to reflect the organism concentration in the mixed culture.

Target organisms for qualitative analyses may also at times be quantified from mixtures by direct inoculation on selective agar media. However, due to stress and competition, this may lead to a lower colony count than the actual concentration in the vials. In such a case the dispersion will also usually become larger.

8.2.3.5. Criteria for homogeneity

Calculations and reasoning by same principle as for drinking water mixtures is made here. This means that the criterion for homogeneity is that the values obtained for the test of reproducibility (T) and the test "Index of dispersion" between vials (I_2) should not exceed the set limits simultaneously, i.e. 2.6 for T and 2 for I_2.

For qualitative analyses, the same criteria prevail and target organisms must be detected from all the 10 vials containing the mixed cultures.

For production of test items before 2015, the criteria of homogeneity for quantitative analyses according to Peterz, 1992 (12) was used. The range for the *log_{10} transformed* values of the mean result of the 10 vials should not exceed 0.5 $\log_{10}$ units and the standard deviation should be <0.15 $\log_{10}$ units. In the original test of Peterz results from only 5 vials were used.

A value of 2 for T was desirable when T was decided to be used as a measure in the Netherlands (17), but in practice a value of T = 3 were used. Here the value T = 2.6 is chosen for food analyses because it corresponds exactly to the standard deviation of 0.15 $\log_{10}$ units that was previously used. For drinking water analyses the desired value T = 2.0 is used based on empirical checks. This difference is justified by the fact that for food analyses more divergent high values are obtained (in theory) that become normalized after $\log_{10}$ transformations in comparison with results from drinking water analyses which are normalized by square-roots.

8.2.4. Measurement uncertainty for the final determination of concentration and homogeneity

When a manufactured material is assessed as homogenous, in that sense that the dispersion is not significantly larger between vials than within, the individual analyses are regarded as 10×2 (20) independent analyses. The pooled standard deviation from the ANOVA is reckoned as the measurement uncertainty for the determination of the concentration that was performed under repeatability conditions during the check of homogeneity.

When making these determinations, transformed analytical results are used just as in the ANOVAs, implying that the *square root transformed* values are used for drinking water analyses and the *common logarithms (log_{10})* for food analyses.

8.3. Stability

Every freeze-dried mixture of microorganisms is not tested for stability. However, initial concentration and homogeneity tests are usually performed with several weeks in between, allowing the detection of unexpected changes in concentrations (e.g. due to loss of vacuum in the vials). But altogether, we rely on many years of experience of freeze-dried material and on continuous follow-ups of the reference materials (RM), constituted of similar microbial mixtures, manufactured and used for internal control.

For proficiency tests, if there has been manufactured more vials of a certain mixture than for one single testing round, a new combined stability check of concentration and homogeneity is carried out if used more than 6 months after the last check. Five randomly selected vials are primarily used, as in a stability test for RM. If the results reveal that the mixture is not homogeneous for one or more analyses, these analyses must be done again on 10 new vials. If the homogeneity is still not accepted, the mixture is discarded.

Stability test for RM are carried out on a regular basis during at least 2 years for material that was kept at −55 °C (before 2012 −65 °C) and, in many cases, at −20 °C. Analyses have been performed after about 1, 6, 12 18 and 24 months after production. In addition, if there has been any doubt, the material has been further analysed. A follow-up is often also made in stabile mixtures as well after about 36 and 48 months.

The analyses correspond to half a homogeneity check, i.e. the content of 5 vials is analysed in duplicates (two equal volumes or two dilution series for drinking water and food, respectively). The statistical analyses correspond to those in the homogeneity tests but are based on the 5×2 results.

In order to detect changes better, the results are plotted over time for each analytical parameter and reference material for the temperature it was kept at.

8.4. Outliers in a testing round

Outliers are results that differ so much from the other results that they cannot be explained by the ordinary variation. Outliers can be objectively identified in different ways. In both the Food and Drinking water schemes, the Grubbs test modified by Kelly (15, 16) is used. The level of 1 % is set as risk to erroneously identify a result as an outlier. A prerequisite for a correct test is that the results are normally distributed.

In order to start with a normal distribution as good as possible, the results are always transformed before being processed. Square root transformation is used for the drinking water results and $\log_{10}$ transformation is used for the food results. The test is used as an objective tool to identify deviating results even when the transformed results are not fully normal distributed. However, in this case the assumption of normal distribution is violated.

Outliers are excluded before the final calculations of medians, mean values and measures of dispersion for the various analyses. However, “z-scores” (see 8.6.9)

are calculated also for the outliers, using the same mean value and standard deviation for a parameter as for ordinary z-scores.

8.5. False positive and false negative results in a testing round

8.5.1. The number of false results

The number of reported false results varies a lot depending on which analysis is performed, the composition of the sample and the degree of difficulty, e.g. concentration and/or background flora.

8.5.2. False positives – definition

A false positive result is an analytical result where an organism is considered detected even though it was not present in the sample.

8.5.3. False negatives – definition

A false negative result is an analytical result where the target organism is not detected in a proper volume even though it should be present in that volume.

8.5.3.1. Additional considerations for the drinking water scheme

When the average colony concentration is high, a result of zero is obviously false and no outlier test is necessary to confirm it.

When the average colony concentration is low, a zero result that is identified as an outlier is defined as a false negative result. A zero result that is not identified as an outlier gets a remark as "False negative ?" (underlined) in the annex with all results in the final report when the average for the parameter is >15 CFU.

When the average colony concentration is very low (e.g. <10 CFU), zero results are often obtained by random and are not identified as outliers, and thus even not as false negative results.

8.6. Statistics for a testing round

8.6.1. Transformations and account for different measures

Calculations of mean values, standard deviations and z-scores are carried out log_{10} *transformed* or *square root transformed* for food and drinking water analyses, respectively (see 8.1).

For drinking water, the mean and the median values are shown in the normal CFU scale (back transformed) in reports, while common logarithmic results are stated for the food scheme.

The various measures in this chapter which are based on a normal distribution (all except median and range) are not accounted for if there is less than a total of 20 results for an analysis. For special analyses with few participants, these measures can be reported anyway but this will be noted in the report.

Median and range are given when the number of results is lower than 20 for an analysis. In this case, false results are reported but not outliers.

Median and range are also reported for parameters that are not used for assessment or giving z scores but used for discussion only. Neither false results nor outliers are reported for these results.

When results are grouped based on the method used, at least 5 results are needed to calculate a measure of dispersion and mean.

8.6.2. Median

Medians are given in the preliminary results of the participants instead of mean values. They are also given in parallel to the mean values in the final report. The median is more robust than the mean value, which means it is less affected by outliers and the distribution.

8.6.3. Range

The range is the interval that includes the lowest to the highest result for a parameter. Depending on whether the analysis is assessed or not, the interval is used after or before the outliers and false results are excluded, respectively. When these deviating results are excluded, the range can be called an *interval of acceptance*.

8.6.4. Mean value

The mean value is calculated from the results of the participating laboratories when outliers and false results have been excluded.

8.6.5. Assigned value (m)

For evaluated parameters, the assigned value is the mean value (8.6.4). It is regarded as the true, normative value. In some cases different assigned values may be relevant, based on groupings of the data according to the method used. For parameters that are not statistically evaluated, the median is stated as the assigned value.

In the reports from the drinking water scheme the re-transformed (squared) assigned value is given as the mean value rather than the assigned value itself.

There are a couple of reasons for choosing the consensus value as the assigned value, rather than a value determined by expert laboratories.

a. Microbiological quantitative results are strongly dependent on method and therefore no “true value” strictly exists. The true value is to some extent a question of definition. Results obtained with a certain method by one or several expert laboratories are not necessarily more correct than those obtained with another method by a participating laboratory.

b. Different brands or batches of dehydrated culturing media manufactured according to a relevant method standard may give different colony appearances and recovery. A systematically specific recovery (“bias”) depending on the culturing medium should not be allowed to have an impact on the outcome for participants using another medium.

8.6.6. Standard deviation for proficiency assessment (s)

This variation around the participants mean value of an analysis is estimated from the actual variation of their results and is made up by the standard deviation after outliers and false results have been excluded. This standard deviation is also used as denominator in the calculation of z-scores (see 8.6.9).

It is also stated as a measure of dispersion in the Food scheme since it *is a relative measure* (independent on concentration) *when logarithms are used*.

An alternative to this standard deviation, which usually varies from round to round for a parameter, would be to have a fixed standard deviation for each parameter. One way is to use an average value based on previous suitable testing rounds. With such a procedure – often recommended – z-sores are considered to be directly comparable from one round to another. In order for this to be relevant, it is presupposed that the test materials of different testing rounds have the same degree of difficulty. If not, a difficult mixture with a large dispersion among the results would lead to a larger number of "extreme" z-scores (z-scores beyond the acceptance limits). However this difficulty is then caused by the manufacturer of the test material and can be considered as something that should not affect an individual laboratory outcome.

The degree of difficulty to obtain a correct result varies with the testing material, both within and between testing rounds – even for the same analytical parameter. The National Food Agency has therefore chosen to use a standard deviation that varies with the difficulty of the testing material, i.e. the actual standard deviation for the current material. In general, this leads to about the same proportion of "extreme" z-scores for the results within accepted limits (i.e. without outliers) in every testing round. Added to this are a varying number of extreme "z-scores" for the outliers present (results outside the accepted limits that were not included in calculations of mean value and standard deviation). The total number of extreme z-scores may therefore vary more or less independently between rounds.

In this way the z-scores are more accurate for the performance over time as well, since a compensation for the degree of difficulty is made. Use of the participant standard deviations is therefore considered to be the most appropriate way to make comparisons of the microbiological analyses within the frame of the program.

8.6.7. Coefficient of variation (CV)

The coefficient of variation (CV) is a *relative measure* and is the standard deviation (8.6.6) in percent of the mean value.

The coefficient of variation is stated as the measure for dispersion in the Drinking water scheme.

8.6.8. Measurement uncertainty for the assigned value

The uncertainty of measurement for an assigned value is calculated as the standard deviation from the testing round divided by the square root of the number of correct results (''standard error'').

The higher the number of reported results that are included in the calculations, the less influence of the uncertainty of measurement compared to the standard deviation for calculation of the z-scores (see 8.6.6). Already with as few as 10 results, the measurement uncertainty makes up only about 30% of the standard deviation and decreases to less than 10% with more than 100 results.

Accuracy that is made up by trueness (bias) and precision (dispersion) of the results is not stated. Differences in recovery due to usage of different methods and/or media are a problematic component of microbiological analyses. This means that systematic deviations from a "true" value cannot be quantified and does not have to be erroneous. The accuracy is therefore only specified as precision and expressed by the measurement uncertainty.

8.6.9. Comparative values for follow-up (z-scores)

All results (outliers included) except the false results from assessed parameters are transformed into standard values (z-scores) according to the formula:

$$z = \frac{x - m}{s}$$

x = the result of the individual laboratory (in transformed form)
m = the mean of the participating laboratories (the assigned value)
s = the standard deviation of the laboratories results around m

After this transformation, the standard values have, apart from those based on outliers, a mean value equal to zero (0) and a standard deviation equal to one (1), and makes up a distribution that can be compared to a standardised normal distribution. The z-scores make it possible to compare the various analyses with each other since they are independent of concentration and expressed according to the same scale (number of standard deviations).

In 95% of the cases, an individual z-score, apart from those based on outliers, will be found in the interval [–2; +2]. The probability to fall outside these limits is less than 5%. The probability to fall outside the interval [–3; +3] is less than 0.3%.

The z-scores that are calculated based on the outliers can be regarded as artificial, not fully true, since the outliers are not included in the calculations of the common mean value and the common standard deviation for an analysis. These "artificial" z-scores exist in addition to the ordinary ones and are in general found outside of the interval [–3; +3].

The z-scores are the base of the box plots (see 8.7.2.4).

8.6.10. Interpretation of the z-scores

In evaluations of the analytical results, the following guidelines may be used:

$\|z\| \leq 2$	indicates that (the original) result is acceptable
$2 < \|z\| \leq 3$	indicates a warning that the result may be deviating, and might possibly motivate an action in the follow-up process
$\|z\| > 3$	indicates that the result is regarded as deviating and should lead to an action in the follow-up process.

8.7. Reports from a testing round

8.7.1. Preliminary report on the website

About 1 week after the last date for registration of the results, a preliminary report is published in two parts on the participant pages of the website (see chapter 2.2.3). The first part is a feed-back on the results registered by the laboratory along with preliminary processed results of all participants with median values and a preliminary relative measure of dispersion. The preliminary limits of acceptance are available for at least 2 weeks. As these preliminary limits of acceptance may change, they are not always shown during processing of the results. During this period the laboratories can only see their own results and the medians, until the final statistics are ready and the final report is published.

The second part is a pdf document with comments to open from the same webpage. The document should be read alongside with the results and contains a short description of the mixtures content and the way to interpret the results outcome with respect to the preliminary limits of acceptance. The report often contains an annex with photos showing outcome and appearance of the colonies on various selective media.

8.7.2. Final report

8.7.2.1. General

A final report is published within 2 months after last date for registration of the analytical results. The report contains discussions of the results of the different parameters analysed but also describes the outcome and performances in some general ways. Deviating results are paid extra attention as well as cases where an analytical parameter was impossible to evaluate in a certain mixture. The results for each relevant parameter are presented in histograms (see 8.7.2.3). In addition, the outcome of results with respect to different methods used is compiled in tables. Possible discrepancies between methods are discussed. An annexed table presents all results obtained by all participants (with identification code), in which outliers and false results are highlighted. Statistics are summarised at the end of this table. Another annex provides the calculated z-scores, intended for individual follow-up of the laboratories. These z-scores are also graphically summarised in box plots, one per laboratory, in order to visualize overall performance (see 8.7.2.4). The report is often rounded off with an annex presenting photos illustrating the outcome on various selective media.

The report is available as a pdf-document, on the participants' pages of the program website (see 2.2.3) and in connection to the information about the proficiency testing program on the public website of the National Food Agency (*www.livsmedelsverket.se*).

8.7.2.2. Quality control results

For both schemes, results obtained from the final determination of concentration and homogeneity test (see 8.2.2.3 and 8.2.3.3) or the last stability/homogeneity check (see 8.3) are presented in a table in the final report.

8.7.2.3. Histograms

A histogram is made for each analysis where quantitative non-zero results are present and illustrates the distribution of the results with special markings for the outliers and false negative values. When relevant, a corresponding histogram is shown with the results differentiated according to different methods or media. The histograms for drinking water are based on original colony counts and the

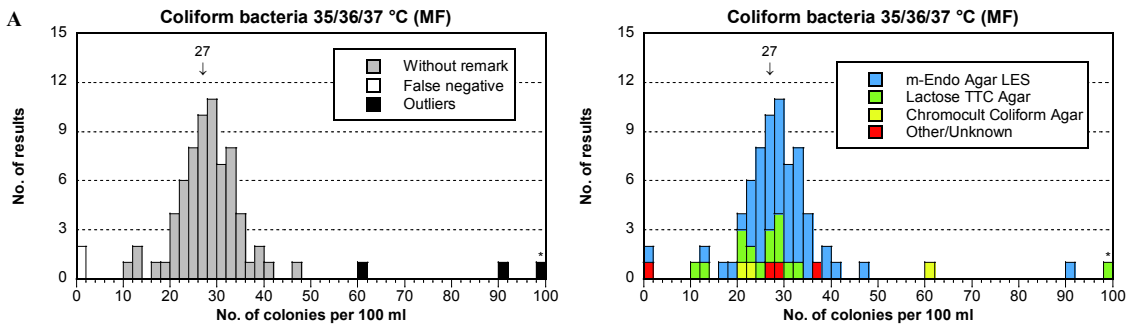


Figure 1 *Example of a histogram for a drinking water analysis*

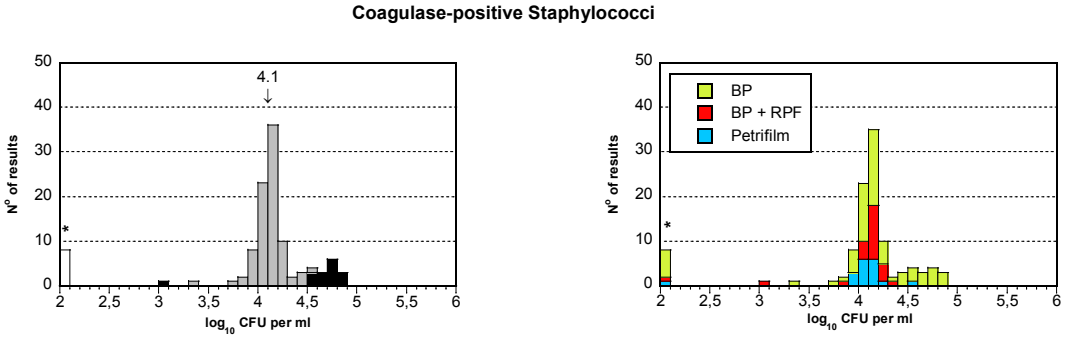


Figure 2 *Example of a histogram for a food analysis*

histograms for food are based on $\log_{10}$ results. Examples from final reports are given in figures 1 and 2. An asterisk indicates that the results are outside the interval of the horizontal axis.

8.7.2.4. Box plots

Each box plot is based on the z-scores of an individual laboratory and illustrates how its standardised results are situated as a group and as a median value in relation to the common, "true", mean value zero. The median value of the laboratory is illustrated by a solid horizontal line in the box. The box is made up by the 50% middle results. The remaining upper and lower 25% of the results are represented by vertical lines and circles. A z-score is represented by a circle when it is deviating by a certain factor in relation to the extent of the box.

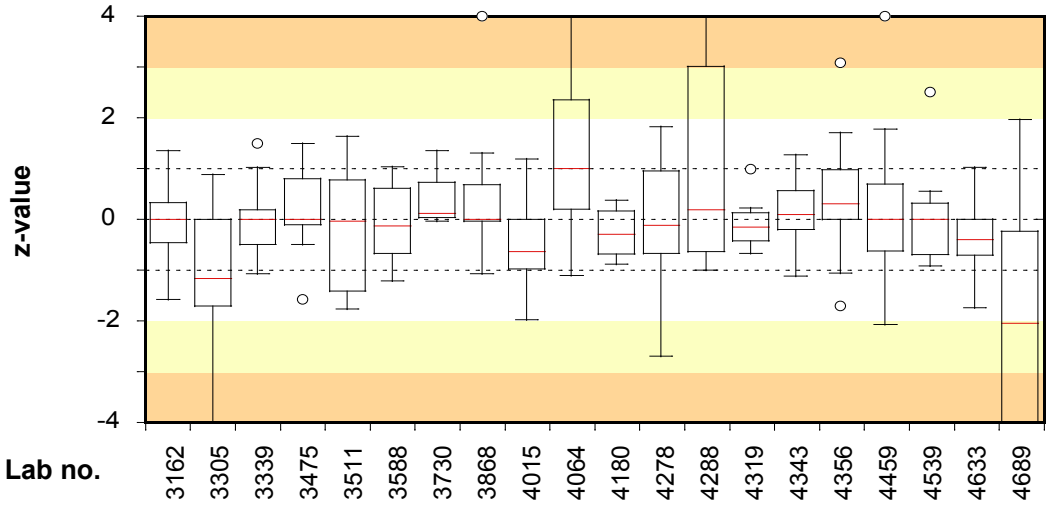


Figure 3 *Example of box plots*

Z-scores higher than +4 or lower than –4 are represented as the limits +4 and –4 of the y-axis in the plots, respectively. An example of a box plot is presented in figure 3.

8.8. Assessing the performance of an individual laboratory

The performance of the individual laboratory is not directly assessed in the final report. The basis for assessment is however provided. The criteria on which the assessment could be based are foremost the number of false negative and false positive results together with the number of outliers (see 8.4 and 8.5) or the numbers of deviating z scores. For each laboratory, the z scores are shown in the box plot and the other numbers are stated in a table below it (see 8.7.2.4). The corresponding results are highlighted (bold and shaded area) in the annex of the final report containing the results of all laboratories, making them easy to connect to a specific laboratory. In the same annex, sample numbers (1, 2, 3 etc.) are crossed over if, based on many analytical results, it is obvious, that samples have been mixed-up.

There is no summary measurement of the overall performance for the testing round and the laboratories are not ranked based on their performance. It is thus up to each participating laboratory in their follow-up procedure to interpret the outcome and its performance. As help, all z-scores (see 8.6.9) of the laboratory are reported in an annex and are further visualized in a box plot (see 8.7.2.4). Outliers are also reflected by high positive (+) or negative (–) z-scores, while the false results do not generate z-scores at all. Z-scores are otherwise mainly intended to compare the results of different analytical parameters in one round and the result of a specific parameter from round to round.

The box plot combined with the numbers of false results and outliers in the table below the plot is the easiest way for an accreditation body, or other interested party, to grasp the performance of a laboratory in a testing round.

9. Confidentiality and user identity

9.1. Confidential laboratory number

Each laboratory is given a unique laboratory number when registering. The number is valid for the schemes the laboratory is registered for and is confidential. It means that the provider reveals the number to no one but the laboratory in question. The provider never states the laboratory number to a third party.

9.2. Confidential password

In addition to the laboratory number, the laboratory is also given a confidential password when registering.

9.3. Usage of laboratory number and password

The laboratory should state its laboratory number when communicating with the provider regarding a scheme. This number is also used by the provider to identify the laboratory on the website, in compilations and in the final reports. Correspondence with a participant where its identity and laboratory number is revealed is confidential.

The laboratory number and the password are used together to log in to the participant pages of the website (see 2.2.3).

The laboratory number and password are stated on a document label enclosed with the test material and the instructions for each round.

9.4. Changing the laboratory number and password

The laboratory number may be changed in order to minimise the risk of unwarranted usage, e.g. upon staff turnover. The password and the laboratory number may also be changed upon participants written request, or if either part has used them in such a way that the identity of the laboratory is violated.

10. Filing

All results that are registered in the test rounds of the two schemes are filed in our database for as long as the laboratory is registered there. Documents derived from correspondence or that have been generated by the results are filed for at least 4 years.

11. Comments from participants and remarks

11.1. Policy

Remarks in the form of deviations and complaints on the work performed within the proficiency testing schemes are documented and investigated. If required, corrective actions and measures to avoid re-occurrence will be taken. Comments and suggestions for improvements are taken into consideration and dealt with in an appropriate way.

11.2. Remarks

Remarks within the proficiency testing schemes are divided in complaints and deviations. A complaint is pertaining when a participant or other interested party is dissatisfied with a service or a product. Deviation applies when written procedures are not followed or when requirements or agreements are not fulfilled. These definitions and the rules for managing the complaints and deviations are

described in general terms in a general instruction (18) and more specifically in the internal procedure of the program.

Complaints and deviations as well as the measures taken are generally documented in the overarching system for these tasks at the National Food Agency. If the issue is more general it will also be managed within the case and document management system at NFA.

11.3. Suggestions for improvement

When there is a possibility to address suggestions for improvement – within which preventive actions are included – they are also documented. The rules for managing these suggestions are described in general term in a general instruction (18) and more specifically in the internal procedure of the program.

12. Conditions and obligations

The general conditions for participation and the obligations of the participating laboratories and the organiser are stated on o website *(see the back side of the front cover)*. Contracts with special conditions and obligations can be established between the organizer and an individual laboratory when necessary.

12.1. Extract of the general conditions of participation

12.1.1. Who can participate?

- Laboratories performing analyses within the frames of the schemes and that are using relevant methods.
- Laboratories to which consignments will be available in time by use of ordinary mailing facilities or by use of courier paying an extra fee and that are able to report results and pay invoices in due time.
- Laboratories that have access to the Internet and are prepared to use the website of the program.

12.1.2. Which methods may be used?

- Methods aimed for the analysis that is performed. The methods should, preferentially, be used as routine methods.

12.1.3. Fee

- An invoice is sent for the testing round(s) the laboratory has registered for.
- The fee has to be paid within the time period stipulated (normally 30 days after print-out of the invoice).

12.2. Other obligations of participating laboratories

- To visit the program website and actively consider participation in the testing rounds.
- To report results according to given instructions.

12.3. Obligations of the National Food Agency

- To keep the information on the website up-to-date regarding testing rounds, analyses, dates and prices.
- To provide the original and preliminary processed results on the website within the stated period of time.
- To publish a final report as a pdf-document on the website within the stated period of time.

12.4. Limited responsibility

- The provider has no liability regarding third party claims depending on a laboratory's participation and performance in any of the schemes run by the National Food Agency.

13. Participation costs

Current prices for the respective schemes are stated on the website. We reserve the right to change the prices if necessary, in order to be able to continue the activity based on the set requirements.

The fee for participation is paid after invoice. The prices are stated in Swedish crowns (SEK) as well as USD ($) and Euro (€) and payment can be made in any of these currencies.

14. This protocol

This protocol will be revised when important modifications are made. Participating laboratories will be notified whenever a new edition is available for download or printing.

15. References

1. EN ISO/IEC 17025:2005. General requirements for the competence of testing and calibration laboratories.
2. EN ISO/IEC 17043:2010. Conformity assessment – General requirements for proficiency testing (ISO/IEC 17043:2010, IDT)

3. Peterz, M. & Steneryd, A.C. 1993. Freeze-dried mixed cultures as reference samples in quantitative and qualitative microbiological examinations of food. J. Appl. Bacteriol. 74:143-148.
4. International Standard, ISO 8199:2005. Water quality – General guidance on the enumeration of micro-organisms by culture. 2nd ed., 2005-06-15.
5. Ordinances of Swedish Work Environment Authority (*Arbetsmiljöverkets författningssamling*), 2005. Microbiological risks of working surroundings – infections, toxin effect, hypersensitivity. AFS 2005:1. *In Swedish*.
6. Kallings, I. & Ljungdahl Ståhle, E. Swedish Institute for Infectious Disease Control, 2002. Expert's report regarding proficiency testing and reference samples from the Swedish National Food Administration. Swedish Institute for Infectious Disease Control Reg. nr. 527/2002-18. National Food Agency Reg. nr. 2509/02. *English translation.*
7. Frändberg, E. & Anér, G. National Food Agency (*Livsmedelsverket*), 2002. Judgement on microbiological material for controls. National Food Agency Reg. nr. 2509/02. *In Swedish*.
8. Thompson, M. & Wood, R. 1993. The international harmonized protocol for the proficiency testing of (chemical) analytical laboratories. Pure & Appl. Chem., Vol. 65, No. 9 2123-2144. *(Also published 1993 in: J. AOAC International 76, 926-940)*
9. Central Science Laboratory, 1997. Food analysis Performance Assessment Scheme (FAPAS®) – Organisation and analysis of data. FAPAS Secretariat, CSL – Food Science Laboratory, Norwich Research Park Conley, NORFOLK NR4 /UQ, United Kingdom. 5th edition, April 1997.
10. Thompson, M., Ellison, S. & Wood, R. 2006. The international harmonized protocol for the proficiency testing of analytical chemistry laboratories. Pure & Appl. Chem., Vol. 78, 145-196.
11. Food analysis Performance Assessment Scheme (FAPAS®), 2002. – Protocol for the Organisation and Analysis of Data. Central Science Laboratory, Sand Hutton, York, YO41 1LZ / United Kingdom. 6th edition, September 2002.
12. Peterz, M. 1992. Proficiency testing: external evaluation of analytical competence in Microbiological Food laboratories. SLV rapport nr 21, 1992, 15 p. *In Swedish*.
13. BCR Information, Chemical analysis.1993. Statistical analysis of certification trials for microbiological reference material. Commission of the European Communities, EUR 15008 EN.
14. Elliott, J. M. 1977. Some methods for statistical analysis of samples of benthic invertebrates. Freshwater Biological Association, Scientific publication no. 25, 2nd edition, 4th impression 1993.
15. Grubbs, F. & Beck, G. 1972. Extension of sample sizes and percentage points for significance tests of outlying observations. Technometrics 14:847-854.

16. Kelly, P. 1990. Outlier detection in collaborative studies. J. Assoc. Off. Anal. Chem. 73:58.64.
17. Mooijman, K. M., During, M. & Nagelkerke, N. J. D. 2003. MICROCRM: Preparation and control of batches of microbiological materials consisting of capsules. RIVM report 250935001/2003. RIVM, Bilthoven, Holland.
18. National Food Agency (*Livsmedelsverket*), The Quality system, Science Division. "Actions in relation to discrepancies, complaints and proposals for improvement". Latest version in the quality manual on the Intranet of NFA. *In Swedish.*
19. ISO 13528:2016 Statistical methods for use in proficiency testing by interlaboratory comparison.

Proficiency Testing

SWEDAC ACKREDITERING
1457
ISO/IEC 17043

Livsmedelsverket
National Food Agency

NFA PT
Since 1981